Project Ideas

Wednesday, October 3, 2018

8:53 AM

Text prediction & auto correct

Three features:

1. Real time spell checking
   1. Teem -> team or tea or teams
   2. If we have time
2. Past spelling correction
   1. Teames -> teams
   2. If we have time
3. Word prediction
   1. Machine learning
   2. What is your -> email, address, name, …
   3. Add sentences they use to dataset
   4. Main focus
4. Punctuation prediction
   1. If we have time

Difficulties

* Data set
* 5,000 words in English language
* Part of speech (adjective, verb)
* Link connected words
  + Quick, quickly

When we pitch it to Bro Burton we could pitch the first one, and then if we have time we can add on 2 and 3

* Start with one, then add on others
  + Word prediction
  + Add punctuation if we feel we can handle it

To find a word bank

* Look at Wikipedia pages
* Download pages
* Unique words
* <https://www.kaggle.com/mikeortman/wikipedia-sentences>

Expectations on how it's run:

* GUI
  + Interested in if we have the time
* Just run with Python

List of priorities:

1. Text prediction
2. GUI
3. Punctuation?
4. Real time spell check
5. Past time spell check

To do:

* Find data sets or word banks
* Explore them
* Brainstorm what needs to be done
  + Keep in mind what you would want to do
* Make a timeline (Jordan)
* Starting next week
  + Have a timeline
  + Delegate portions

Apriori algorithm